МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2007
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2007
Proceedings of the Thirty Sixth Spring Conference of
the Union of Bulgarian Mathematicians
St. Konstantin & Elena resort, Varna, April 2–6, 2007

A REVIEW OF RISK PREDICTORS FOR CENTRAL VENOUS CATHETER RELATED SEPSIS AND DEATH IN CRITICALLY ILL PATIENTS*

Krasimira St. Prodanova

In the study the incidence of and risk factors for Central Venous Catheter (CVC) related sepsis and lethal outcome are investigated based on the data from the patients of the Department of Anaestheology and Intensive Care Medicine of Military Medical Academy, Sofia. A multiple logistic regression is performed to obtained an adjusted estimate of the odds ratios and to identify which factors were associated idependently with these outcomes.

1. Introduction. Central Venous Catheters (CVCs) permit hemodynamic monitoring and allow access for the administration of fluids, blood products, medications, and total parental nutration. With possible exeptions of pulmonary artery catheters, central venous catheters (CVCs) have the highest reported rates of infection and sepsis of all intravascular catheters [1, 2]. Critically ill patients who develop bloodstream infections and sepsis are at a greater risk of death than patients with comparable severity of illness without this complication. A number of factors may contribute to the risk of catheter related infections and sepsis in intensive care patients [2, 3, 4]. A survey of CVCs inserted into patients in the Department of Anaestheology and Intensive Care Medicine of Military Medical Academy, Sofia is presented. A multiple logistic regression is performed in order to identify the risk factors of CVCs related infections/sepsis and lethal outcome. The software package used for statistical modeling of real data was STATISTICA 6.0.

2. Measurments. During the study period 118 CVCs in female and male patients are prospectivly studied. All 118 catheters are inserted in intensive care units by expirienced anaesthetists under strict aseptic techniques. For the insertions two cases could ocure: planed and urgent. Two types of catheters are used – "Selinger" and "Cavafix" with 1 or 2 lumens. Data obtained for each catheter included the patients' APACE (Aqute Phistology Chronic Health Evaluations) score and primary diagnoses on admission (categorized as trauma, postsurgical and other) on the first catheter day. Data are also obtained on patients age, genus, insertion site, clinical and laboratory data pertaining to infections, antibiotic administration, duration of catheterization.

---

[*] **2000 Mathematics Subject Classification:** 62H15, 62P10.
Key words: logistic regression, application to medical sciences.

3. Statistical Model. The regression model is used to relate a categorical response (dependent variable $\boldsymbol{Y}$) to the explanatory variables (predictors) $\boldsymbol{x}_i$ $(i = 1, \dots, c)$. We are interested in the responces: $\boldsymbol{Y_1}$ – CVC related infections and sepsis or not and $\boldsymbol{Y_2}$ – Lethal outcome or not.

About the Dichotomous Responce Models: Let us define the dummy random variable to indicate the two cathegories by $Y = 1$ for cathegory A and $Y = 0$ for B. The probability density for Y given the parameter p is therefore point binomial

$$f(Y/p) = p^Y (1-p)^{1-Y}.$$

We assume that the probability p depends on a linear function

$$d(x_1, \dots, x_c) = \beta_0 + \sum_{i=1}^{c} \beta_i x_i, \tag{1}$$

where x_i $(i = 1, \dots, c)$ are the explanatory variables (the independent variables), β_i are constants. So that, the joint conditional density is

$$f(y_1, \dots, y_n/p(d_1), \dots, p(d_n)) = \prod_{j=1}^{n} [p(d_j)]^{y_j} [1 - p(d_j)]^{(1-y_j)},$$

where n is the size of a random sample of data y for response variable Y.

To be able to relate value of y to the value of d, a most specific assumptions about the form of $p(d)$ is required. In the so called logit or logistic model [5], the distribution function of logistic density is:

$$p(d) = \frac{e^d}{1 + e^d}.$$

The shape of $p(d)$ (logistic distribution) is quite similar to the shape for normal distribution. The odds or odds ratio of a dichotomous response is given by

$$\text{odds ratio} = \left[\frac{p(d)}{1 - p(d)}\right]. \tag{4}$$

The logit transformation of logistic distribution function

$$\ln\left[\frac{p(d)}{1 - p(d)}\right] = d = \beta_0 + \sum_{i=1}^{c} \beta_i x_i \tag{5}$$

gives an important advantage of the model because (5) is a linear function of the explanatory variables. The Newton-Raphson iterative procedure is usually used to make maximal likelihood estimator $\widehat{\beta}$ of the coefficient vector β in the logistic model. The procedure is based on a preliminary estimator of β given by

$$\widehat{\beta} = (X^T X)^{-1} X^T Y,$$

where Y is the vector of y_i response values $(i = 1, \dots, n)$ and $X(n \times c)$ is the matrix of observations. The maximum likelihood is obtained by solving the system of $c + 1$ equations:

$$\sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} y_i x_i, \tag{6}$$

where

$$p_i = \exp\left(x_i^T \beta\right) / (1 + \exp\left(x_i^T \beta\right)). \tag{7}$$

The solutions of these equations given by $\widehat{\beta}$ can be used to obtain the estimator $\widehat{p}_i$ for each of n observations and, hence, the fitted sum $\sum\limits_{i=1}^{n} \widehat{p}_i x_i$ is equal to the observed sum in the right-hand side of (6). In comparison to the multiple linear regression model, the coefficient vector $\widehat{\beta}$ must be interpreted differently:

- The coefficients $\widehat{\beta}$ were interpreted as estimates of log odds ratios.
- A marginal one unit increase in x_j brings an increase in d (i.e. in log odds ratios) of the amount of $\widehat{\beta}_j$.
- The confidence intervals were calculated for the odds ratio estimates by taking the exponent of upper and lower endpoints of the asymptotic confidence interval for the log odds ratio.

Testing of hypothesis concerning the regression parameters can include test of single parameter, test involving several parameters from the same regression, and joint tests involving parameters from different regressions. In polychotomous logistic regression, tests for contribution of one or more parameters from the same regression are usually constructed with a large sample Wald test, with test statistic

$$Q_W = \hat{\beta}^T [\mathrm{Var}(\widehat{\beta})]^{-1} \widehat{\beta},$$

where $\mathrm{Var}(\widehat{\beta})$ is the estimated covariance sub-matrix for the relevant parameters. This statistics is approximately distributed as a $\chi^2(r)$ random variable with r degrees of freedom under the null hypothesis that r-dimensional vector $\widehat{\beta}$ is equal to $\mathbf{0}$. When there is a single parameter of interest, then the test statistic is

$$Q_W = [\widehat{\beta}_j / \mathrm{SE}(\widehat{\beta}_j)]^2 \quad ,$$

where $\mathrm{SE}(\widehat{\beta}_j)$ is the standard error of $\widehat{\beta}_j$ and its distribution is χ^2 $(r = 1)$. Testing the parameters from different regressions (inferences regarding the coefficients) can be made by comparing a full model with c explanatory variables plus an intercept to a reduced model with $(c - r)$ explanatory variables plus an intercept using a likelihood ratio test. The statistic

$$\text{Final loss} = -2(\ln L_c - \ln L_{c-r}) \tag{10}$$

has a chi-square distribution with (r) degree of freedom, where L_c is the likelihood function for the full model, and L_{c-r} is the likelihood function for the reduced model. The likelihood ratio $\chi^2(c-r)$ for the significance of all c variables is given by (10), where $r = 0$.

4. Results and conclusion. Catheter related infections and sepsis were suspected

Table 1. Infections and Sepsis

Model: Logistic regression (logit) N of 0's:56 1's:62

Dep. var: VAR1 Loss: Max likelihood (MS-err. scaled to 1)

Final loss: 72.817019982 ChiI(10) = 17.643 **p** = .06133

	Const. B0	VAR2	VAR3	VAR5	**VAR6**	VAR7	VAR8	**VAR9**	VAR10	VAR11	VAR12
Estimate	−2.4349	0.011	−0.517	0.0341	0.55	−0.105	0.802	1.009	0.036	−0.718	0.0638
Standard Error	0.97678	0.011	0.4407	0.0332	0.21	0.451	0.59	0.374	0.034	0.6198	0.4668
t(107)	−2.4928	0.991	−1.173	1.0261	2.62	−0.232	1.359	2.695	1.073	−1.158	0.1368
p-level	0.01421	0.324	0.2432	0.3071	0.01	0.817	0.177	0.008	0.286	0.2495	0.8915
−95%CL	−4.3712	−0.011	−1.391	−0.032	0.13	−1	−0.37	0.267	−0.031	−1.946	−0.862
+95%CL	−0.4985	0.033	0.3565	0.0999	0.96	0.79	1.973	1.751	0.103	0.511	0.9893
Wald's Chi-square	6.21389	0.982	1.377	1.053	6.86	0.054	1.847	7.261	1.152	1.3406	0.0187
p-level	**0.01268**	**0.322**	**0.2406**	**0.3048**	**0.01**	**0.816**	**0.174**	**0.007**	**0.283**	**0.2469**	**0.8912**
Odds ratio	0.08761	1.011	0.5962	1.0347	1.73	0.901	2.231	2.743	1.037	0.4879	1.0659
−95%CL	0.01264	0.989	0.2489	0.9687	1.14	0.368	0.692	1.306	0.97	0.1428	0.4225
+95%CL	0.60742	1.033	1.4284	1.1051	2.61	2.204	7.192	5.762	1.109	1.667	2.6892
Odds ratio (range)		1.992	0.5962	2.5099	5.14	0.901	2.231	7.524	4.421	0.4879	1.0659
−95%CL		0.502	0.2489	0.4242	1.49	0.368	0.692	1.705	0.284	0.1428	0.4225
+95%CL		7.91	1.4284	14.851	17.8	2.204	7.192	33.21	68.88	1.667	2.6892

Table 2. Infections and Sepsis

Classification of Cases (dterz1∼1.sta)

Odds ratio: 3.0175

		Pred.	Pred.	Percent
		0	1	Correct
Observ.	0	32	24	57.1
Observ.	1	19	43	69.4

in 53% of the total number of 118 lines. The cases with letal outcome were 48, i.e. about 41% of the patients. The responces are: $\boldsymbol{Y_1}$ – CVC related infections and sepsis or not, and $\boldsymbol{Y_2}$ – Lethal outcome or not.

The dummy coding is used to construct categorical variables.

- CVC related infections and sepsis – "1", not – "0" (VAR1);
- Lethal outcome – "1", not – "0" (VAR4).

The explanatory variables in the model for $\boldsymbol{Y_1}$ are 10:

- The age: min = 18, max = 80 (VAR2);
- The gender: male – "0", female – "1" (VAR 3);
- APACHE score: min = 2, max = 29 (Var5);
- The diagnostic group: nonsurgical – "0", trauma – "1", st. surgical – "2", neurosurgical – "3" (VAR6);
- The used types of antibiotics are: more than two –"1", less or two –"0" (VAR7);
- The type of CVC: Selinger – "0", Cavafix – "1" (VAR8);
- The insertion site: v. Femoralis – "−1", v. Subclavia – "0", v. Jugularis – "1" (VAR9);
- The duaration of catherization (days): min = 4, max = 45 (VAR10);
- The number of CVC-lumens: one – "1", more than one – "0" (VAR11);
- Type of catetherizion: planed – "0", urgent – "1".

The explanatory variables in the model for $\boldsymbol{Y_2}$ are 11 – the above 10 plus $\boldsymbol{Y_1}$.

We use the logit model to study the impact of the explanatory variables on the response variable. The software package STATISTICA 6.0 was used for statistical modeling of the data. The results for the response $\boldsymbol{Y_1}$ – CVC related infections and sepsis or not are presented in the Table 1. The likelihood ratio $\chi^2(c-r)$ statistics with $c-r=10$ for the significance of all 10 variables is 72.817, which have p-value 0.06. The fitted logistic regression model indicates that at the margin the probability that CVC will relate infection or sepsis decrease with decreases in age, APACHE scores and duration of catheterization. The values of Wald's chi-square test using 0.05 level of significance suggest that CVC related infections or sepsis is related to diagnostic group ($p = 0.009$) and the insertion side ($p = 0.007$). The remaining variables are insignificant. The fitted model was used to obtain value of $\widehat{p}_i$ ($i = 1, \ldots, n$) from (7) for each observation. If $\widehat{p}_i < 0.5$, the observation is placed in the category $Y_1 = 0$, and otherwise – in $Y_1 = 1$. The prediction success matrix based on the $\widehat{p}_i$ is given in the Table 2. From 56 cases 32 are correctly coded "0" and 43 of the 62 are corectly coded "1".

Table 3. Lethal outcome

Model: Logistic regression (logit) N of 0's:69 1's:49

Dep. var: VAR4 Loss: Max likelihood (MS-err. Scaled to 1)

Final loss: 45.005716095 ChiI(11) = 70.165 **p** = .00000

B0	VAR1	VAR2	**VAR3**	**VAR5**	VAR6	VAR7	VAR8	VAR9	**VAR10**	**VAR11**	VAR12	
Estimate	−6.4533	0.6	0.016	−1.915	0.29	0.382	−0.89	−0.56	0.054	0.084	1.607	−1.106
Standard Error	1.5662	0.583	0.015	0.654	0.058	0.278	0.6	0.7157	0.532	0.043	0.79	0.7018
t(106)	−4.1203	1.03	1.037	−2.929	4.997	1.376	−1.48	−0.782	0.102	1.965	2.035	−1.577
p-level	7.5E–05	0.305	0.302	0.004	2E–06	0.172	0.14	0.4361	0.919	0.052	0.044	0.1179
$-95\%CL$	−9.5584	−0.56	−0.01	−3.211	0.175	−0.17	−2.08	−1.979	−1	−7E–04	0.042	−2.498
$+95\%CL$	−3.3481	1.756	0.046	−0.619	0.405	0.933	0.3	0.8595	1.11	0.169	3.173	0.285
Wald's Chi-square	16.9772	1.061	1.075	8.58	24.97	1.894	2.19	0.6112	0.01	3.863	4.143	2.4855
p-level	**3.8E–05**	**0.303**	**0.3**	**0.003**	**6E–07**	**0.169**	**0.14**	**0.4343**	**0.919**	**0.049**	**0.042**	**0.1149**
Odds ratio (unit ch)	0.00158	1.823	1.016	0.147	1.336	1.466	0.41	0.5715	1.056	1.088	4.99	0.3307
−95%CL	7.1E–05	0.574	0.986	0.04	1.191	0.845	0.13	0.1383	0.368	0.999	1.043	0.0823
+95%CL	0.03515	5.791	1.047	0.539	1.499	2.543	1.35	2.3619	3.033	1.184	23.89	1.3297
Odds ratio (range)		1.823	2.704	0.147	2492	3.149	0.41	0.5715	1.115	33.92	4.99	0.3307
−95%CL		0.574	0.404	0.04	111.9	0.603	0.13	0.1383	0.135	0.97	1.043	0.0823
+95%CL		5.791	18.11	0.539	55490	16.44	1.35	2.3619	9.199	1187	23.89	1.3297

Table 4. Lethal outcome

Classification of Cases (dterz1∼1.sta)

Odds ratio: 18.462

		Pred.	Pred.	Percent
		0	1	Correct
Observ.	0	60	9	86.96
Observ.	1	13	36	73.47

The results for the response $\boldsymbol{Y_2}$ – Lethal outcome or not are presented in the Table 3. The likelihood ratio $\chi^2(c-r)$ statistics with $c-r=11$ for the significance of all 11 variables ($r=1$ in (10)) is 45.00, which have p-value 0.000. The fitt is, therefore, perfect. The fitted logistic regression model indicates again that at the margin the probability of mortality decrease with decreases in age, APACHE scores and days of cathetarization. For the dummy variables, a insertion site v. Jugularis is more likely to get lethal outcome, whereas if the genus of a patient is mail, it is less likely. The values of Wald's chi-square test using 0.05 level of significance, suggest that lethal outcome is related to genus ($p=0.003$), APACHE score ($p=0.000$), the duration of catheterization ($p=0.049$) and the number of CVC lumens ($p=0.042$). The remaining variables are insignificant. The prediction classification matrix based on the $\widehat{p}_i$ is given in the Table 4. On the basis of the results of this study the rate of CVC related basing practices on knowledge of risk factors could reduce infections or sepsis and lethal outcome.

REFERENCES

[1] M. L. Pearson. Guideline for prevention of intravascular device-related infections. Part II. Recommendations for the prevention of nosocomial intraveascular device-related infections. Hosp. Infect. Control Pract. Advisory Committee. *Am. J. Infect Control*, **24** (1996), 277–293.

[2] N. P. O'Grady. Guideline for prevention of intravascular catheter-related infections. Catheter for Disease Control and Prevention. *MMWR Recomm. Rep.* **51** (2002), 1–29.

[3] E. Tacconelli et al. CVC-Related Sepsis in a Cohort of 366 Hospitalised Patients. *Eur. J. Clin. Microbiol. Infect. Dis.*, **16** (1997), 203–209.

[4] S. Oncu et al. CVC Related Infections: Risk Factors and the Effect of Glycopeptide Antibiotics. *Annals of Clinical Microbiology and Antimicrobials*, **2** (2003), 1–6.

[5] J. D. Jobson. Applied Multivariate Data Analysis, Vol. II. Springer Verlag, 1991.

Krasimira Stojanova Prodanova
Technical University of Sofia
Faculty of Applied Mathematics and Informatics
1156 Sofia, Bulgaria
e-mail: kprod@tu-sofia.bg

ОБЗОР НА РИСКОВИТЕ ФАКТОРИ ЗА ЦЕНТРАЛНИТЕ ВЕНОЗНИ КАТЕТРИ, СВЪРЗАНИ СЪС СЕПСИС И СМЪРТ ПРИ КРИТИЧНО БОЛНИ

Красимира Ст. Проданова

В работата са изследвани рисковите фактори за централни венозни катетри (ЦВК), свързани със сепсис и летален изход въз основа на данни от пациенти на Отделението по анестезиология и реанимация на Военно-Медицинската Академия в София. Използвана е многомерна логистична регресия за намирането на адекватни оценки за съотношенията на шансовете и идинтифициране на факторите, които се свързват с тези изходи.